Biochem. J. (1986) 240, 69–74 (Printed in Great Britain)

Cat pancreatic eicosapeptide and its biosynthetic intermediate

Conservation of a monobasic processing site

Henrik Vendelbo NIELSEN, Ulrik GETHER and Thue W. SCHWARTZ*
Laboratory of Molecular Endocrinology, University Department of Clinical Chemistry, Rigshospitalet 6321, DK-2100 Copenhagen, Denmark

Pancreatic eicosapeptide is synthesized together with the hormone pancreatic polypeptide in a common precursor in the major endocrine cell type of the duodenal pancreas. This processing has been previously demonstrated in man and in the dog. In the present study the cat pancreatic eicosapeptide and a *C*-terminally extended form of this were isolated and characterized from acid/ethanol extracts of pancreas by gel filtration and reverse-phase h.p.l.c. The sequence homology in the *C*-terminal part of the eicosapeptides from different species was shown to continue to the other side of the monobasic cleavage site in the extended intermediate form, whereas the end of the extension differed both in chain length and amino acid sequence. It is concluded that the processing site in the intermediate form of the pancreatic eicosapeptide is an example of a proline-directed monobasic cleavage site that has been conserved during evolution.

INTRODUCTION

The pancreatic polypeptide family of regulatory peptides currently consists of three homologous peptides: pancreatic polypeptide (PP), neuropeptide Y (NPY) and peptide YY (PYY). All three peptides consist of 36 amino acid residues with a *C*-terminal tyrosine (Y) amide residue. NPY is an important neuropeptide in the central and peripheral nervous systems. It is often located and functioning together with noradrenaline (norepinephrine) (Tatemoto, 1982*a*). PYY is a hormone located predominantly in the distal small intestine (Tatemoto, 1982*b*). PP, the oldest recognized member of the family, is a hormone released from the major endocrine cell type of the duodenal pancreas and regulates pancreatic exocrine secretion and biliary tract motility (Schwartz, 1983). PP is co-synthesized with another peptide product, pancreatic eicosapeptide (20 amino acids), on a common precursor (Schwartz *et al.*, 1980; Schwartz & Tager, 1981). The precursor is relatively small, comprising only an additional oligopeptide of five to seven amino acid residues in addition to a signal peptide of 29 amino acid residues, which gives a total of 95 amino acid residues (Boel *et al.*, 1984; Leiter *et al.*, 1984; Takeuchi & Yamada, 1985). The precursor for NPY is homologous to the PP precursor with a *C*-terminal extension of similar size but not obviously homologous in primary structure to the eicosapeptide part of the pancreatic polypeptide precursor (Minth *et al.*, 1984).

Processing of precursors for regulatory peptides generally occurs at dibasic sequences, as originally outlined by Steiner *et al.* (1974). However, within recent years it has become evident that single basic residues can also serve as processing sites (Schwartz, 1986). The common precursor for PP and pancreatic eicosapeptide is cleaved both at a dibasic site, between the two peptides, and at a monobasic site, after the eicosapeptide, as studied in detail in dog endocrine cells (Schwartz *et al.*, 1980; Schwartz & Tager, 1981; T. W. Schwartz, unpublished work). In order to determine to what degree the monobasic cleavage site has been conserved during evolution, in the present study we have isolated and characterized the pancreatic eicosapeptide and its biosynthetic intermediate form from cat pancreas.

MATERIALS AND METHODS

Peptide isolation

The duodenal part of cat pancreas was excised within 30 min *post mortem*. The tissue (60 g from nine cats) was frozen on solid CO_2 and stored at −80 °C. It was broken into small pieces and homogenized in a Braun blender in cold acidified ethanol [final concentrations 68% (v/v) ethanol and 0.1 M-HCl) at −20 °C. Homogenization was continued intermittently until the temperature reached 4 °C. The homogenate was incubated overnight at 4 °C and then centrifuged at 10000 *g* for 30 min at the same temperature in a Sorval RC-5B High Speed centrifuge with a GSA rotor. The supernatant was neutralized with 25% (v/v) NH_3 and centrifuged at 5000 *g* for 5 min at 4 °C in the same rotor and centrifuge. Proteins in the new supernatant were precipitated by adding 2 vol. of ethanol and 4 vol. of diethyl ether. The precipitate that formed during incubation overnight at 4 °C was dried under N_2 and reconstituted in 10 ml of 1 M-acetic acid. The extract was gel-filtered on a 2.5 cm × 95 cm column of Sephadex G-50 (superfine grade) (Pharmacia, Sweden) eluted with 1 M-acetic acid at 4 °C at a flow of 28 ml/h. Absorption at 280 nm was determined in the 7 ml fractions, and samples from selected fractions were analysed on an h.p.l.c. system calibrated with insulin, PP and dog pancreatic eicosapeptide (for details see below). Fractions were pooled, and the solvent was evaporated under vacuum. The peptides were reconstituted in 0.5 ml of 3 M-acetic acid and finally purified by h.p.l.c. on a 0.8 cm × 25 cm reverse-phase column of Techogel C_{18} (5 μm particles, 30 nm pore diameter) (packing material purchased from HPLC-

Abbreviations used: PP, pancreatic polypeptide; NPY, neuropeptide Y.

* To whom correspondence should be addressed.

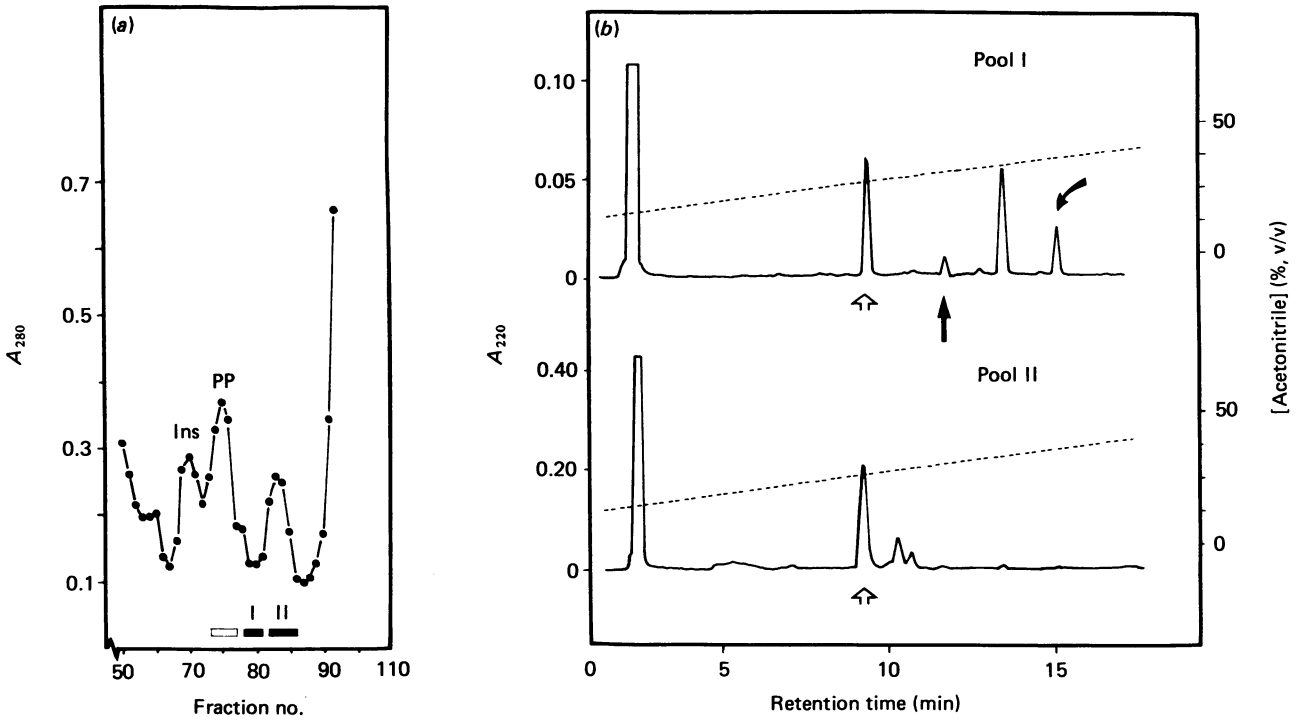

Fig. 1. Purification of peptides from acid/ethanol extraction and ether precipitation of cat duodenal pancreas

(*a*) Gel-filtration profile of peptides on a Sephadex G-50 (superfine grade) column. Only peptides of an apparent M_r from approx. 8000 to 1000 are shown. Fractions were pooled according to the indicated bars. The elution positions of insulin (Ins) and pancreatic polypeptide (PP) are indicated. (*b*) H.p.l.c. profile of peptides from pool I and pool II of the gel filtration. The peptide marked with an open arrow was shown (sequence determination) to be the pancreatic eicosapeptide, and the peptide indicated by a solid black arrow was the *C*-terminally extended form of the eicosapeptide (the intermediate form). The curved arrow points to the elution position of cat PP.

Technology, Macclesfield, Cheshire, U.K.). The chromatography was performed on a Hewlett–Packard 1090 liquid chromatograph with 0.1% (v/v) trifluoroacetic acid (Merck, Darmstadt, Germany) in water as the aqueous phase and acetonitrile (Merck) as the organic solvent.

Enzymic digestion

Tryptic fragments of the h.p.l.c.-purified peptides were obtained after removal of the solvent under vacuum and reconstitution of the peptides in 0.2 ml of 0.1 M-Tris buffer (Trizma base from Sigma Chemical Co., St. Louis, MO, U.S.A.) adjusted to pH 8.0 with HCl, containing 0.01 M-$CaCl_2$ and 0.01 mg of L-tosylphenylalanylchloromethane-treated trypsin (Worthington Corp., Freehold, NJ, U.S.A.)/ml. The digestion was performed at 37 °C for 30 min and terminated by heating the solution for 1 min at 100 °C. Tryptic fragments were purified by h.p.l.c. as described above.

Digestion of h.p.l.c.-purified peptides with carboxypeptidase Y (generously given by J. Johansen, Carlbiotech, Copenhagen, Denmark) was performed at 25 °C. Solvent was removed under vacuum, and the purified peptides were reconstituted in 0.2 ml of *N*-ethylmorpholine/acetate buffer [0.1 M-*N*-ethylmorpholine (Sequanal grade; Pierce Chemical Co., Rockford, IL, U.S.A.) adjusted to pH 7.0 with acetic acid]. Carboxypeptidase Y (3.5 μg in 10 μl of water) and as internal standard 5 nmol of *p*-fluorophenylalanine dissolved in 5 μl of 50% (v/v) methanol (Merck) containing 1% (v/v) triethylamine (Sequanal grade; Pierce Chemical Co.) was added, and 25 μl samples were collected after 0, 0.5, 1, 3, 5, 10 and 30 min of digestion and heated for 1 min at 100 °C. Free amino acids were analysed as phenylthiocarbamoyl derivatives by reverse-phase h.p.l.c. on a Merck Supersphere C_{18} column (0.4 cm × 25 cm) after modification with phenyl isothiocyanate (Sequanal grade; Pierce Chemical Co.) by the method of Heinriksson & Meredith (1984).

Sequence determination

Solvent was removed under vacuum from the h.p.l.c. purified peptides. They were reconstituted in 0.06 ml of 0.1% (v/v) acetic acid, and subjected to automated sequence analysis by sequential Edman degradation on an Applied Biosystems 470A gas-phase Sequenator with the MHTFA1 program of M. Hunkapiller [a program modified from Hunkapiller *et al.* (1983), available at Applied Biosystems, Foster City, CA, U.S.A.] and heptane as a supplementary solvent S_1 (all chemicals were purchased from Applied Biosystems). The phenylthiohydantoin derivatives of amino acids were characterized by h.p.l.c. on a Hewlett–Packard 1084 liquid chromatograph with a 0.45 cm × 25 cm column of CN [(5 μm particles (IBM Instruments, White Plains, NY, U.S.A.)] and a sodium acetate/acetonitrile gradient elution system as described previously (Hunkapiller & Hood, 1983). The samples from the Sequenator were methylated before h.p.l.c. by treating the dried derivatives with acidified methanol (1 M-HCl in methanol;

Applied Biosystems) for 10 min at 50 °C. Aminobutyric acid was used as internal standard during the h.p.l.c. for correction of elution time and for quantifying the amino acid derivatives.

RESULTS

Analytical h.p.l.c. of selected fractions from the gel filtration of peptides, obtained by acid/ethanol extraction and ether precipitation of cat duodenal pancreas, indicated that a peptide similar to the dog pancreatic eicosapeptide was eluted in a peak distinct from the PP peak during gel filtration (Fig. 1*a*). Fractions corresponding to this peak, fractions 82–86 (II), and those immediately before, fractions 78–81 (I), were pooled. As demonstrated in Fig. 1, pool II contained only one major peptide, indicated by the open arrow, which by sequence determination was shown to be homologous to the previously characterized pancreatic eicosapeptides. This peptide could also be identified in pool I, in which a minor peptide peak (indicated by the closed black arrow) by sequence determination was shown to have the same *N*-terminal sequence but to be extended by seven amino acid residues in its *C*-terminal end. This peptide is homologous to the biosynthetic intermediate form of the eicosapeptide characterized in dog endocrine cells and has been designated the intermediate form. Since some of the amino acid residues in the *C*-terminal end of the intermediate form could not be determined with certainty during the Edman degradation of the intact molecule, the *C*-terminal extension on the eicosapeptide sequence was obtained by h.p.l.c. of tryptic fragments of the intermediate form (Fig. 2). Tryptic fragments of both the eicosapeptide and the intermediate form were purified and the single fragment unique to the intermediate form was sequenced (indicated by arrow in Fig. 2). This approach could be used because the eicosapeptide terminates with an arginine residue. The yields of the phenylthiohydantoin derivatives of amino acids obtained in the different sequence cycles are shown in Table 1. No phenylthiohydantoin derivative was identified in sequence cycle number 8 during the Edman degradation of the tryptic fragment of the intermediate form. Also, no free amino acids were identified during carboxypeptidase treatment of the intermediate form of the eicosapeptide, indicating that no further amino acid residues are found after the Val-27. Neither Asp-26 nor Val-27 was demonstrated in reasonable amounts during the degradation, which is in accordance with the fact that digestion with carboxypeptidase Y is slowed down considerably by a penultimate acidic amino acid residue, especially at the high pH and low temperature that were used in order to limit the degradation to the *C*-terminal end (Hayashi *et al.*, 1975). In parallel experiments the enzyme degraded other peptides (results not shown). Furthermore the sequence is in accordance with the sequence of the human intermediate form previously characterized by deduction from the cDNA sequence (Boel *et al.*, 1984).

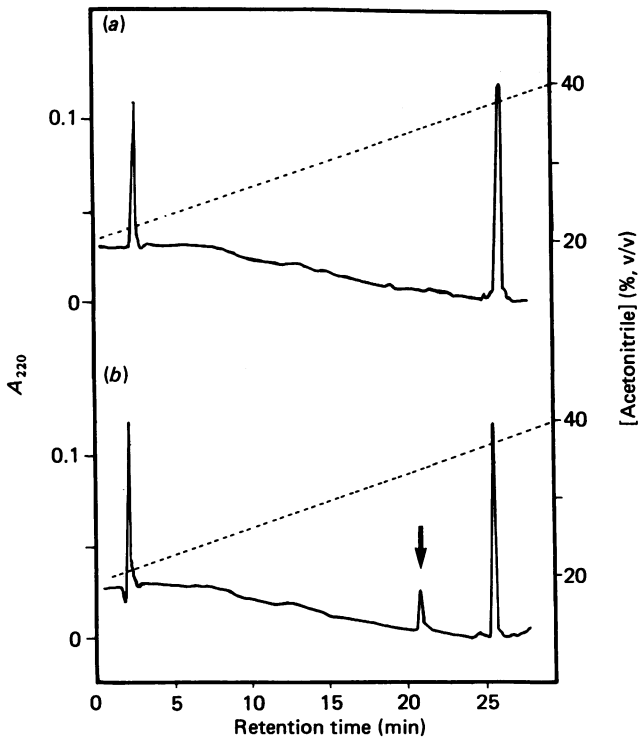

Fig. 2. Isolation by h.p.l.c. of the *C*-terminal extension of the intermediate form of the pancreatic eicosapeptide

Pancreatic eicosapeptide (*a*) and the intermediate form (*b*) were treated with trypsin, and the fragments were purified by h.p.l.c. as in Fig. 1. The arrow indicates the peptide unique to the intermediate form that was used for sequence determination (note that the eicosapeptide terminates in an arginine residue).

Table 1. Sequence determination of peptides from cat pancreatic extract

Sequence cycle	Amino acid	Yield of amino acid phenylthiohydantoin derivative of amino acid (pmol): Eicosapeptide	Intermediate form	*C*-Terminal tryptic fragment
1	Asp	1098	1282	
2	Arg	422	1656	
3	Gly	1832	3360	
4	Glu	1309	2720	
5	Thr	354	965	
6	Leu	1997	1830	
7	Asp	981	1373	
8	Ile	1577	1630	
9	Leu	1940	1325	
10	Glu	1041	914	
11	Trp	364	481	
12	Gly	821	1063	
13	Ser	10	94	
14	Pro	499	618	
15	His	290	306	
16	Ala	667	505	
17	Ala	772	509	
18	Ala	619	355	
19	Pro	362	279	
20	Arg	—	218	
21 (1)	Glu		106	2227
22 (2)	Leu		87	886
23 (3)	Ser		—	150
24 (4)	Pro		116	605
25 (5)	Met		—	1203
26 (6)	Asp		84	366
27 (7)	Val		—	215

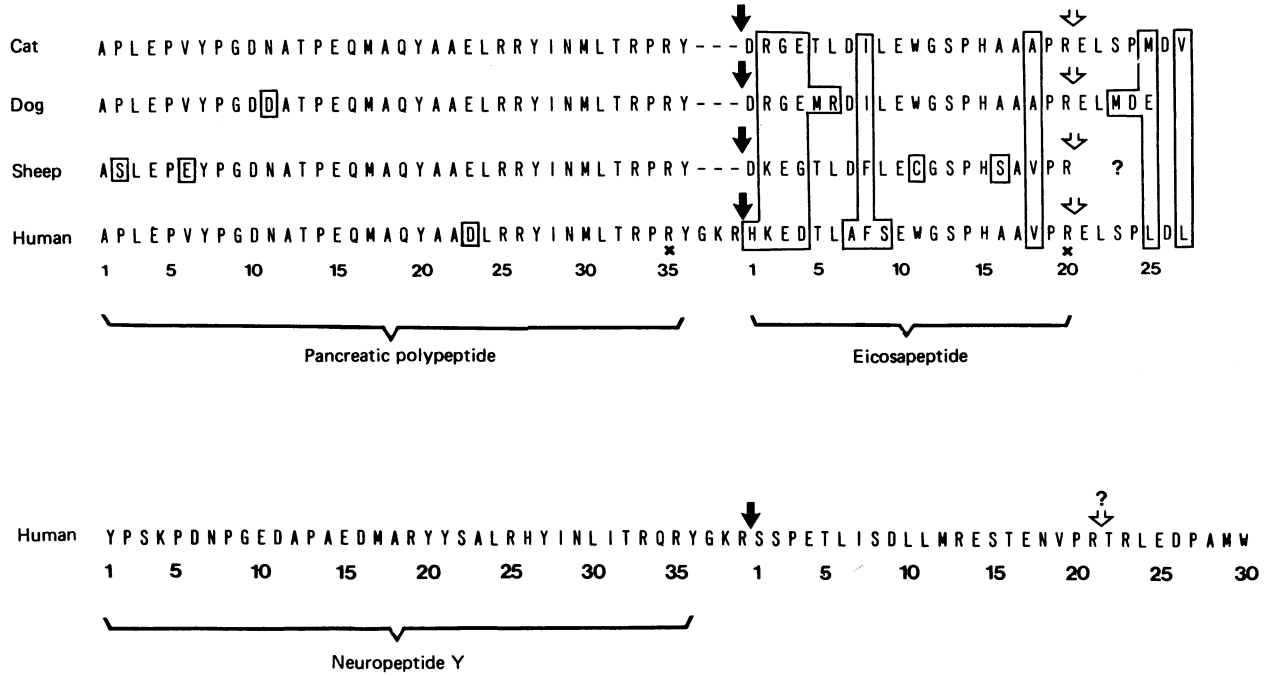

Fig. 3. Comparison of the amino acid sequences of the common precursor for pancreatic polypeptide and pancreatic eicosapeptide in different species

The open arrow indicates the monobasic cleavage site and the solid black arrow indicates the dibasic cleavage site. Asterisks indicate the location of introns in the human gene (Leiter *et al.*, 1985). The structure of the human propeptide is both deduced from a cDNA sequence (Boel *et al.*, 1984) and determined by peptide chemistry characterization of the main peptides (Chance *et al.*, 1979; Schwartz *et al.*, 1984). The question mark in the *C*-terminal end of the sheep precursor indicates the presence of a probable extension. However, only the structure of the eicosapeptide itself has been determined (Schwartz & Hansen, 1984). The PP sequences have all been determined by Chance *et al.* (1979), except for that of the cat, which was performed along with the present study (isolated from the peptide pool indicated with an open bar in Fig. 1; detailed results not presented). Below, the cDNA-deduced sequence of the human neuropeptide Y (NPY) propeptide is shown (Minth *et al.*, 1984), in which the open arrow with the question mark indicates the presence of a suggested 'proline-directed arginine cleavage' site. The one-letter code for amino acids is used.

The sequence of the cat eicosapeptide is very similar to that of the dog counterpart, the only differences being in positions 5 and 6 (Fig. 3). On the other hand, the *C*-terminal extension of the intermediate form varies with respect to sequence and length, i.e. seven amino acid residues in cat and man but only five in the dog. However, the amino acid sequence surrounding the monobasic cleavage site in the intermediate form, -Pro-Arg-Glu-Leu-, is conserved in all species studied (Fig. 3).

DISCUSSION

In the present study the pancreatic eicosapeptide and a *C*-terminally extended form of this peptide were isolated and characterized from cat pancreas. The *C*-terminally extended form is homologous to a biosynthetic intermediate of the eicosapeptide that has been characterized in pulse–chase experiments in isolated endocrine cells from dog pancreas (Schwartz & Tager, 1981; T. W. Schwartz, unpublished work). In previous studies we have shown that the pancreatic eicosapeptide itself is more highly conserved during evolution in its *C*-terminal region than in its *N*-terminal region (Schwartz & Tager, 1981; Schwartz *et al.*, 1984; Schwartz & Hansen, 1984). By comparing the intermediate forms from different species, it appears that conservation of primary structure in the *C*-terminal end continues to the other side of the monobasic processing site. Thereafter distinct differences are observed in chain length and amino acid sequences (Fig. 3). Apparently, the structural basis for the generation of the eicosapeptide from the PP precursor through a monobasic processing mechanism has been conserved. It may be of significance that one of the intervening sequences in the PP gene is located just in this monobasic cleavage site (indicated with an asterisk in Fig. 3). The other intron in the coding sequence is located just before the dibasic processing site (Leiter *et al.*, 1985). The introns thus separate structural domains that eventually will be released as distinct peptides.

The eicosapeptide is a secretory product from the PP cell, but no convincing biological effect has yet been attached to it. It is possible, however, that the monobasic cleavage site at this location in the precursor actually is of greater importance in other members of the PP peptide family, e.g. in the precursor for NPY. Here a similar site for 'proline-directed arginine cleavage' (see below) is found at almost the same place in the homologous precursor although there is little similarity in the primary structure in this part of the precursor (Fig. 3). Thus it could be suggested that the cleavage site could have been conserved in order to ensure the correct production of another peptide also ending in Pro-Arg

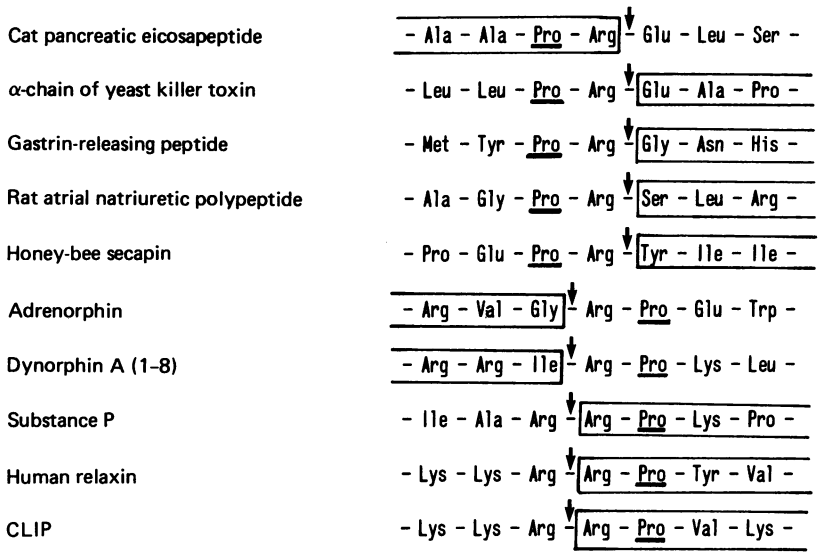

Fig. 4. Monobasic processing sites with adjacent proline residues in precursors for regulatory peptides

The main peptide product is indicated by an open box and its name is shown to the left. Adjacent proline residues are underlined. Arrows indicate the putative initial cleavage site. Information concerning the precursor structure, the specific cleavage site or the peptide are from pancreatic eicosapeptide (the present study), yeast killer toxin (Bostian *et al.*, 1984), gastrin-releasing peptide (Reeve *et al.*, 1983), atrial natriuretic polypeptide (Kangawa *et al.*, 1984), secapin (Vlasak & Kreil, 1984), adrenorphin (Matsuo *et al.*, 1983), dynorphin A (Cone *et al.*, 1983), substance P (Nawa *et al.*, 1983), human relaxin (Hudson *et al.*, 1983) and corticotropin-like intermediate-lobe peptide (CLIP) (Nakanishi *et al.*, 1979).

but co-synthesized and possibly co-functioning with NPY.

The main processing sites for peptide precursors are pairs of basic residues (Steiner *et al.*, 1974). However, as recently reviewed, the rapidly growing knowledge of precursor structure obtained mainly through deduction from cDNA sequences has revealed that cleavage at single basic residues is also fairly common (Schwartz, 1986*a*). In PP cells in culture it is possible to distinguish between the monobasic cleavage mechanism, which occurs late in the biosynthetic process and is lost during the first 24–48 h of culture, and dibasic processing, which occurs early and is retained even after 10 days of culture (T. W. Schwartz, unpublished work). In about one-third of cases of monobasic cleavage, the basic residue is either preceded or followed by a proline residue (Fig. 4). If the sequence is Pro-Arg, the scissile bond is the one on the *C*-terminal side of the arginine residue, but it is on the *N*-terminal side if the sequence is Arg-Pro. The special conformational constraint, induced by proline residues on the peptide backbone, has led to the suggestion that the three-dimensional structure is particularly important for the processing at single basic residues (Schwartz, 1986). As shown in Fig. 4, such 'proline-directed arginine cleavage' is found even in yeast (Bostian *et al.*, 1984), and in the present investigation evidence is presented that a particular monobasic site in the intermediate form of pancreatic eicosapeptide is highly conserved among different mammalian species.

The present study was supported by grants from the NOVO Foundation, the Carlsberg Foundation and the Danish Natural Science Research Council. H.V.N. and U.G. are the recipients of pregraduate scholarships from the Weimann Foundation and T.W.S. of a professorship in molecular endocrinology under the Danish Medical Research Council and the Weimann Foundation. Frede Hansen is thanked for technical assistance and Ms. Henny Jensen for secretarial help during the preparation of the manuscript. We also thank Mairead O'Hare for helpful discussion and comments on the manuscript.

REFERENCES

Boel, E., Schwartz, T. W., Norris, K. E. & Fiil, N. P. (1984) EMBO J. **3**, 909–912

Bostian, K. A., Elliott, Q., Bussey, H., Burn, V., Smith, A. & Tipper, D. J. (1984) Cell **36**, 741–751

Chance, R. E., Moon, N. E. & Johnson, M. G. (1979) in Methods of Hormone Radioimmunoassay, (Jaffe, B. M. & Behlman, H. R., eds.), pp. 667–672, Academic Press, New York

Cone, R. I., Weber, E., Barchas, J. D. & Goldstein, A. (1983) J. Neurosci. **3**, 2146–2152

Hayashi, R., Bai, Y. & Hata, T. (1975) J. Biochem. (Tokyo) **77**, 69–79

Heinriksson, R. L. & Meredith, S. C. (1984) Anal. Biochem. **136**, 65–74

Hudson, P., Haley, J., Cronk, J. M., Crawford, R., Haralambidis, J., Tregear, G., Shine, J. & Niall, H. (1983) Nature (London) **301**, 628–631

Hunkapiller, M. W. & Hood, L. E. (1983) Methods Enzymol. **91**, 486–493

Hunkapiller, M. W., Hewick, R. M., Dreyer, W. J. & Hood, L. E. (1983) Methods Enzymol. **91**, 399–413

Kangawa, K., Tawaragi, Y., Oikawa, S., Mizuno, A., Sakuragawa, Y., Nakazato, H., Fukuda, A., Minamino, N. & Matsuo, H. (1984) Nature (London) **312**, 152–155

Leiter, A. B., Keutmann, H. T. & Goodman, R. H. (1984) J. Biol. Chem. **259**, 14702–14705

Leiter, A. B., Montminy, M. R., Jamieson, E. & Goodman, R. H. (1985) J. Biol. Chem. **260**, 13013–13017

Matsuo, H., Miyata, A. & Mizuno, K. (1983) Nature (London) **305**, 721–723

Minth, C. D., Bloom, S. R., Polak, J. M. & Dixon, J. D. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 4577–4581
Nakanishi, S., Inoue, A., Nakamura, M., Chang, A. C. Y., Cohen, S. N. & Numa, S. (1979) Nature (London) **278**, 423–427
Nawa, H., Hirose, T., Takashima, H., Inayama, S. & Nakanishi, S. (1983) Nature (London) **306**, 32–36
Reeve, J. R., Walsh, J. H., Chew, P., Clark, B., Hawke, D. & Shively, J. E. (1983) J. Biol. Chem. **258**, 5582–5588
Schwartz, T. W. (1983) Gastroenterology **85**, 1411–1425
Schwartz, T. W. (1986) FEBS Lett. **200**, 1–10
Schwartz, T. W. & Hansen, H. F. (1984) FEBS Lett. **168**, 293–298
Schwartz, T. W. & Tager, H. S. (1981) Nature (London) **294**, 589–591
Schwartz, T. W., Gingerich, R. L. & Tager, H. S. (1980) J. Biol. Chem. **255**, 11494–11498
Schwartz, T. W., Hansen, H. F., Håkanson, R., Sundler, F. & Tager, H. S. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 708–712
Steiner, D. F., Kemmler, W., Tager, H. S. & Peterson, J. D. (1974) Fed. Proc. Fed. Am. Soc. Exp. Biol. **33**, 2105–2125
Takeuchi, T. & Yamada, T. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1536–1539
Tatemoto, K. (1982*a*) Proc. Natl. Acad. Sci. U.S.A. **79**, 2514–2518
Tatemoto, K. (1982*b*) Proc. Natl. Acad. Sci. U.S.A. **79**, 5485–5489
Vlasak, R. & Kreil, G. (1984) Eur. J. Biochem. **145**, 279–282

Received 7 March 1986/2 June 1986; accepted 14 July 1986